Award Number: W81XWH-11-1-0527

TITLE: Collaborative Model for Acceleration of Individualized Therapy of Colon Cancer

PRINCIPAL INVESTIGATOR: Aik Choon Tan

CONTRACTING ORGANIZATION: UNIVERSITY OF COLORADO, DENVER
AURORA, CO 80045-2505

REPORT DATE: December 2015

TYPE OF REPORT: FINAL REPORT

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| December 2015 | Final Report | 15Sep2011 - 14Sep2015 |

**4. TITLE AND SUBTITLE**

Collaborative Model for Acceleration of Individualized Therapy of Colon Cancer

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-11-1-0527

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Aik Choon Tan, Ph.D., S. Gail Eckhardt, M.D. and Todd M. Pitts, M.S.

E-Mail: gail.eckhardt@ucdenver.edu;aikchoon.tan@ucdenver.edu;todd.pitts@ucdenver.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
UNIVERSITY OF COLORADO, DENVER
AURORA, CO 80045-2505

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

Colorectal cancer (CRC) represents a major health burden, and is the third leading cause of cancer deaths in the U.S. In the past decade, the median survival among patients with metastatic CRC has significantly improved, primarily due the development of active chemotherapeutic regimens that include biological agents. However, despite this success, patients soon run out of therapeutic options and receive salvage therapy that results in only a few weeks of disease stability. We have proposed to employ a team science, systems biology based approach to rapidly identify novel anti-cancer agents and individualize therapeutic strategies in preclinical CRC models. In this Final Report, we will present the tasks and key accomplishments achieved within the funding of this project. We have published two papers, and are in the preparation of another three manuscripts from the results generated from this project. In summary, we have accomplished all the tasks in this proposal.

**15. SUBJECT TERMS**
Colorectal cancer, novel anti-cancer agents, bioinformatics, next-generation sequencing, predictive biomarkers, individualized medicine

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | USAMRMC |
| U | U | U | UU | 24 | 19b. TELEPHONE NUMBER *(include area code)* |

## Table of Contents

**FINAL REPORT.**

**Introduction:** Colorectal cancer (CRC) represents a major health burden, and is the third leading cause of cancer deaths in the U.S. In the past decade, the median survival among patients with metastatic CRC has significantly improved, primarily due the development of active chemotherapeutic regimens that include biological agents. However, despite this success, patients soon run out of therapeutic options and receive salvage therapy that results in only a few weeks of disease stability. This is particularly true for a subset of patients that have a mutation in the KRAS gene, since it has been shown that one of these new treatments is not effective for them. Therefore, new agents are needed that can stabilize disease and hopefully prolong life in patients with CRC. One of the lessons learned in CRC, in fact, in patients with the KRAS mutation in their tumor, is the importance of not only developing new effective drugs, but also developing ways to select patients for those treatments. Unfortunately the lack of such strategies is what led to thousands of CRC patients with KRAS mutations being treated with epidermal growth factor receptor (EGFR) inhibitors at considerable toxicity and no benefit, when it was discovered that tumors with this mutation did not respond to these drugs. This new area of patient selection, or individualized therapy, is based upon a robust set of research tools in the field of bioinformatics. Therefore, successful research teams are comprised of clinicians, who treat patients with cancer, and bioinformaticians, that are able to synthesize large sets of data and look for patterns of response or resistance to a particular new drug. Such a team has been assembled for this proposal. Thus, the overall goal of this Idea Award is enhance the efficiency and speed of developing novel and individualized therapy for patients with KRAS mutant colorectal cancer (CRC) using a comprehensive bioinformatics approach and novel preclinical models of human CRC. This proposal has the potential of providing novel, individualized therapeutic strategies for CRC patients with KRAS mutations that are poised for clinical testing at the completion of this work (3 years). The yield will be highly relevant, as new drug development will not only be jump-started by this proposal but agents to be tested clinically will be tailored for specific populations of patients with CRC, thereby potentially conferring greater clinical benefit. We will describe our research achievements and outcomes of this project in this **Final Report**.

**Aim 1.** To develop predictive classifiers for 3 novel agents using preclinical models of colorectal cancer (CRC). We have identified the following three novel agents to develop predictive classifier using preclinical models of CRC and these agents will be tested in **Aim 2**.

**Aim 1.** To develop predictive classifiers for 6 novel agents using preclinical models of colorectal cancer (CRC).

We initially selected the following six novel agents to develop predictive classifiers using preclinical models of CRC and three of these agents were tested in **Aim 2**.

**Table 1:** Three novel anti-cancer agents selected in this study.

| **Agents** | **Targets** | **Company** | **Clinical Developmental Phase** |
|---|---|---|---|
| MLN8237 (alisertib) | Aurora Kinase A (AURKA) | Millennium Pharmaceuticals/Takeda | Phase I/II |
| MLN0128 | TORC1/TORC2 | Millennium Pharmaceuticals/Takeda | Phase I/II |
| TAK733 | Dual specificity mitogen-activated protein kinase kinase 1 (MAP2K1) | Millennium Pharmaceuticals/Takeda | Phase I |
| TAK960 | Polo-like Kinase 1 (PLK1) | Millennium Pharmaceuticals/Takeda | Phase I |
| ENMD2076 | Aurora Kinase A (AURKA) and Angiogenic Kinase (KDR) | CASI Pharmaceuticals | Phase I/II |
| PF-04691502 (PF-502) | Phosphatidylinositol 3-Kinase (PIK3CA) and mammalian Target of Rapamycin (mTOR) | Pfizer | Phase I |

**Task 1: *In vitro* cell line exposure (Months 1-12, Dr. Eckhardt).**
To evaluate the sensitivity of CRC cell lines to MLN8237, ENMD2076, and MLN0128, a panel of CRC cell lines were exposed to increasing concentrations of these novel anti-cancer agents and assessed for proliferation using an SRB or CyQuant assay as previously described (Skehan et al 1990; Pitts et al 2010). As depicted in **Figure 1** there was a broad range of sensitivity of the CRC cell lines to these anti-cancer agents, *indicating that patient selection is needed*. ***See Dr. Eckhardt's Final Report for details.***

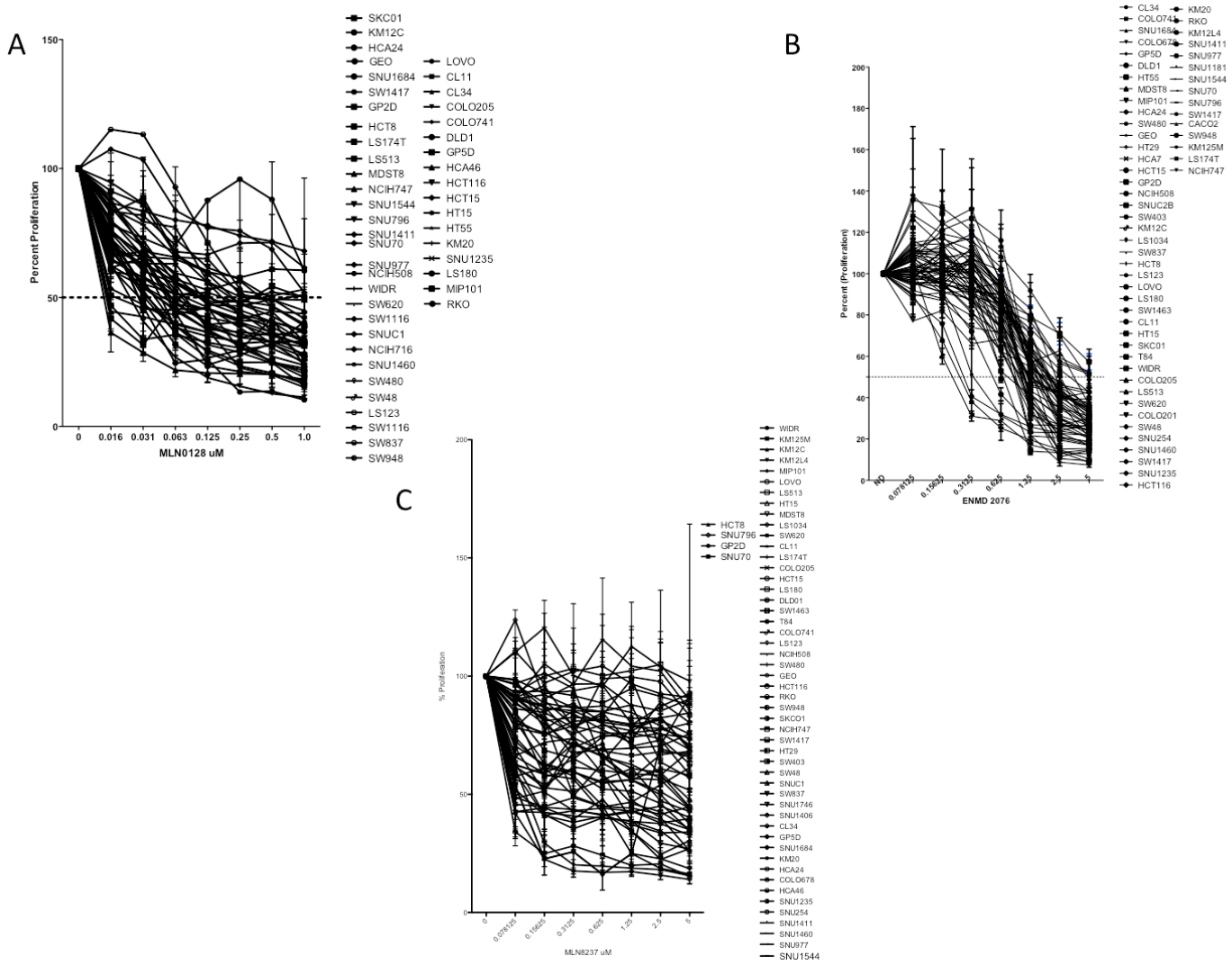


**Figure 1:** A panel of CRC cell lines were exposed to increasing concentrations of MLN0128 (A), ENMD2076 (B), MLN8237 (C).

**Task 2: *In vivo* cell line xenograft treatment (Months 6-18, Dr. Eckhardt).**
To determine the *in vivo* inhibition, we have performed treatment using these anti-cancer agents on cell lines derived xenografts as previously described (Pitts et al 2010). We have treated CRC cell line xenografts with MLN8237 (**Figure 2**), MLN0128 (**Figure 3**), ENMD2076 (**Figure 4**). We are in the process finishing this task by injecting more mice with CRC cell lines and treating with the compounds listed. As anticipated, there is also a diversity of responses to these agents *in vivo.* ***See Dr. Eckhardt's Final Report for details.***

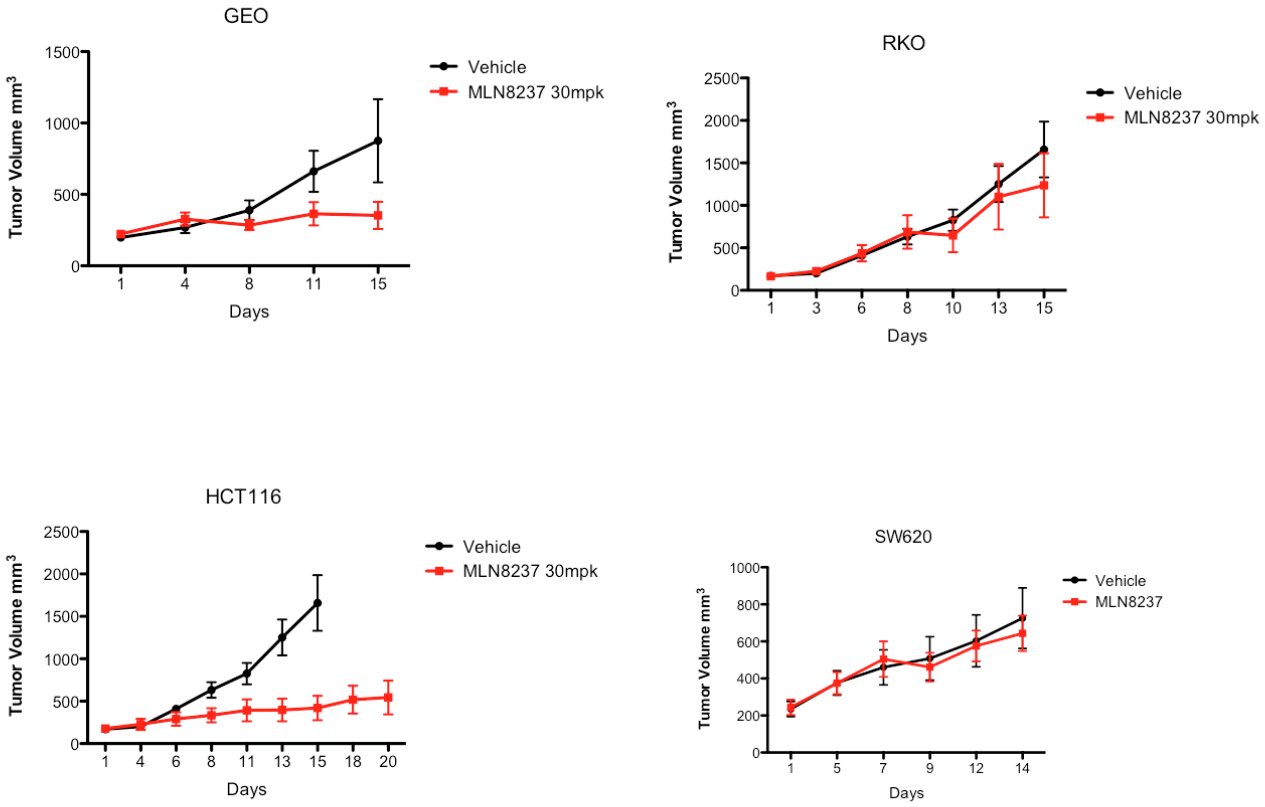


**Figure 2:** *In vivo* cell lines treated with MLN8237.

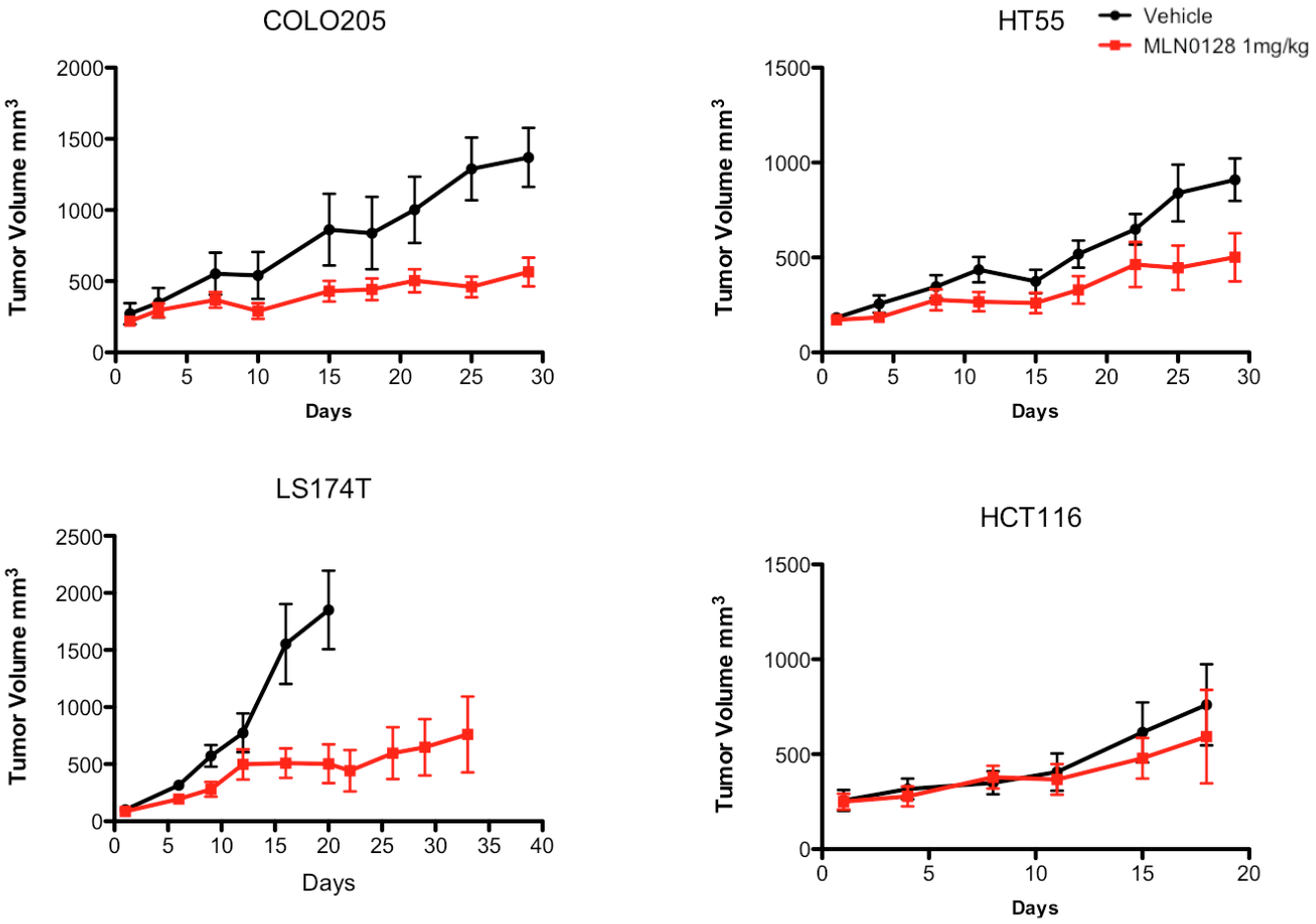


**Figure 3:** *In vivo* cell lines treated with MLN0128.

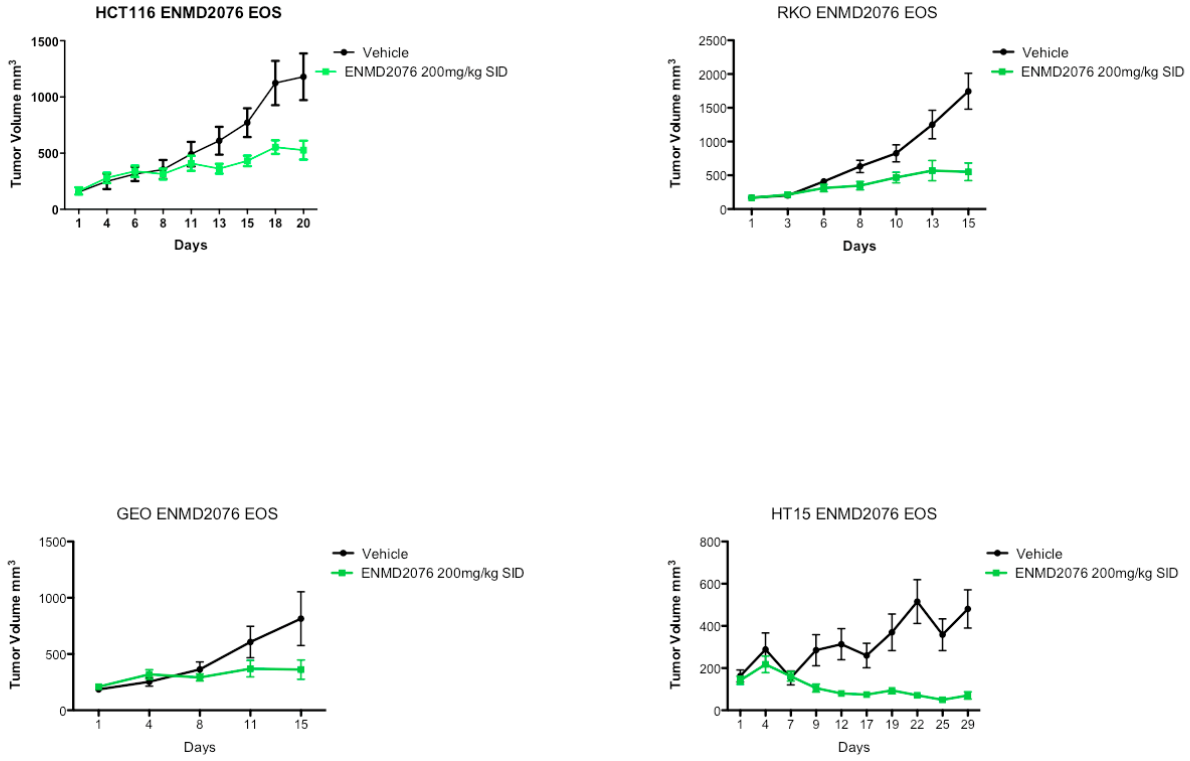


**Figure 4:** *In vivo* cell lines treated with ENMD2076.

**Task 3: Immunoblotting for relevant downstream effectors (Months 6-18, Dr. Eckhardt).**
To access the inhibition of these anti-cancer agents in the cancer cells, we have performed immunoblotting for relevant downstream effectors of these targets. ***See Dr. Eckhardt's Final Report for details*.**

**Task 4: Perform transcriptome sequencing (RNA-Seq) on CRC cell lines (*in vitro* and xenografts) (Months 1-18, Dr. Tan).**
Total RNAs were extracted from the cancer cells or tumor tissues using Trizol (Invitrogen, Carlsbad, CA). Libraries were constructed using 1μg total RNA following Illumina TruSeq RNA Sample Preparation v2 Guide. The poly-A containing mRNA molecules were purified using poly-T oligo-attached magnetic beads. Following purification, the mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were converted into first strand cDNA using reverse transcriptase and random primers. This was followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments then were subjected to an end repair process, the addition of a single "A" base, and ligation of the adapters. The products were purified and enriched using PCR to create the final cDNA library. The cDNA library was validated on the Agilent 2100 Bioanalyzer using DNA-1000 chip. Cluster generation was performed on the Illumina cBot using a Single Read Flow Cell with a Single Read cBot reagent plate (TruSeq SR Cluster Kit v3-cBot-HS). Sequencing of the clustered flow cell was performed on the Illumina HiSeq 2000 using TruSeq SBS v3 reagents. We used the Illumina HiSeq2000 as this is the latest machine with higher sequencing throughput and cheaper for sequencing cost. Utilizing the latest HiSeq2000 machine, we were able to multiplex 3 samples per lane, sequence with single end 100 cycles (1x100bp) and achived ~40 million reads per sample. The number of cycles for each read is also programmed into the machine before the run begins. Sequencing images were generated through the sequencing platform (Illumina HiSeq 2000). The raw data were analyzed in four steps: image analysis, base calling, sequence alignment, and variant analysis and counting. An additional step was required to convert the base call files (.bcl) into *_qseq.txt files. For multiplexed lanes/samples, a de-multiplexing step is performed before the alignment step.

**Task 5: Bioinformatics analysis of RNA-Seq data (Months 12-18, Dr. Tan).**
High-throughput mRNA sequencing (RNAseq) of each sample was obtained from the Illumina HiSeq2000. On average, we obtained about 60 million (coverage ranged from 30 to 90 million reads) single-end 100bp sequencing reads per sample. To analyze the RNAseq data, the reads were mapped against the human genome using the BiNGS! (Bioinformatics for Next Generation Sequencing) pipeline. In our pipeline, we have optimized the parameters for mapping using Tophat (Trapnell et al 2009) and cufflinks (Trapnell et al 2010). The first step of the BiNGS! pipeline is mapping the reads against the reference genome. Here, we used the NCBI reference annotation (build 37.2) as a guide, and allowing 3 mismatches for the initial alignment and 2 mismatches per segment with 25 bp segments using Tophat (version 1.3.2). On average, 92% (ranging from 71% to 95%) and 84% (ranging from 68% to 92%) of the reads aligned to the human genome for cell lines and human CRC explants, respectively. Next, the workflow employed Cufflinks (version 1.3.0) to assemble the transcripts using the RefSeq annotation as the guide, but allowing for novel isoform discovery in each sample. Isoforms were ignored if the number of supporting reads was less than 30 and if the isoform fraction was less than 10% for the gene. The data were fragment bias corrected, multi-read corrected, and normalized by the total number of reads. On average, the sequences can be mapped to 20,221 known genes (ranging from 18,213 to 21,448 genes) and 19,355 known genes (ranging from 17,481 to 21,519 genes) for cell lines and human CRC explants, respectively. The transcript assemblies for each sample were merged using cuffmerge. To estimate the transcript expressions of individual sample, we computed the FPKM values of the transcripts by rerunning Cufflinks again using the merged assembly as the guide. The final output of this analysis step is a P x N matrix, where P is the number of samples and N is the number of transcripts, respectively. Gene expression for individual sample is estimated by summing the FPKM values of multiple transcripts that represent the same gene. Subsequent data analyses of RNAseq will be performed on this matrix. We also performed variants calling analysis on the RNA-seq for all the models using GATK workflow (McKenna et al 2010). We used ANNOVAR (Wang et al 2010) to annotate the functional annotation of these variants. We prioritized on the variants that were predicted as non-synonymous mutations.

**Task 6: Development of the *k*-TSP classifier from mRNA-Seq (Months 18-24, Dr. Tan).**
Using the drug sensitivity data obtained from **Task 1**, we have selected the 5-8 most sensitive (S) and 5-8 most resistant (R) cell lines as the training set for each anti-cancer agent. Using the RNA-seq data from these selected cell lines, we have employed k-TSP algorithm to derive gene pairs as classifier for the selected agent. Internal leave-one-out cross-validation (LOOCV) was performed to avoid overfitting of the training process. On average, these classifiers achieved 75% (range 65% - 85%) of LOOCV accuracies. The number of gene pairs selected in the classifiers was 3 – 9 pairs. The training set for the three selected agents (MLN0128, MLN8237 and ENMD2076) and the gene markers for k-TSP classifier are presented in **Tables 2, 3, and 4** respectively.

**Task 7: Development of an integrated classifier (Months 18-24, Drs. Eckhardt and Tan).**
From the mRNA-seq, we obtained mutations data for the training set cell lines, and initiate evaluation of incorporating KRAS, BRAF, PIK3CA, APC, and TP53 mutations into the k-TSP do not enhance the predictive accuracy of the integrated classifiers. This suggests that the usual suspects of the CRC "driver" genes are not predictive against these novel agents. We have expanded the process of adding additional mutations and/or selected genes within a pathway to refine the predictive accuracy of the integrated classifiers. We also incorporated recent published data (e.g. Diamond et al 2013) into the integrated classifier such as relevant genes described in other cancer types in this refinement process. However, we do not see improvement in the prediction accuracy in the selected three agents.

**Task 8: Prioritization of agents to progress to Specific Aim 2 (Months 18-24, Drs. Eckhardt and Tan).**
We have identified the following three anti-cancer compounds to move into Aim 2:

| Agents | Targets | Company | Clinical Developmental Phase |
|---|---|---|---|
| MLN0128 | TORC1/TORC2 | Millennium Pharmaceuticals/Takeda | Phase I |
| MLN8237 (alisertib) | Aurora Kinase A (AURKA) | Millennium Pharmaceuticals/Takeda | Phase I/II |
| ENMD2076 | Aurora Kinase A (AURKA) and Angiogenic Kinase (KDR) | EntreMed | Phase I/II |

**MLN0128 Training Set and k-TSP classifier.**

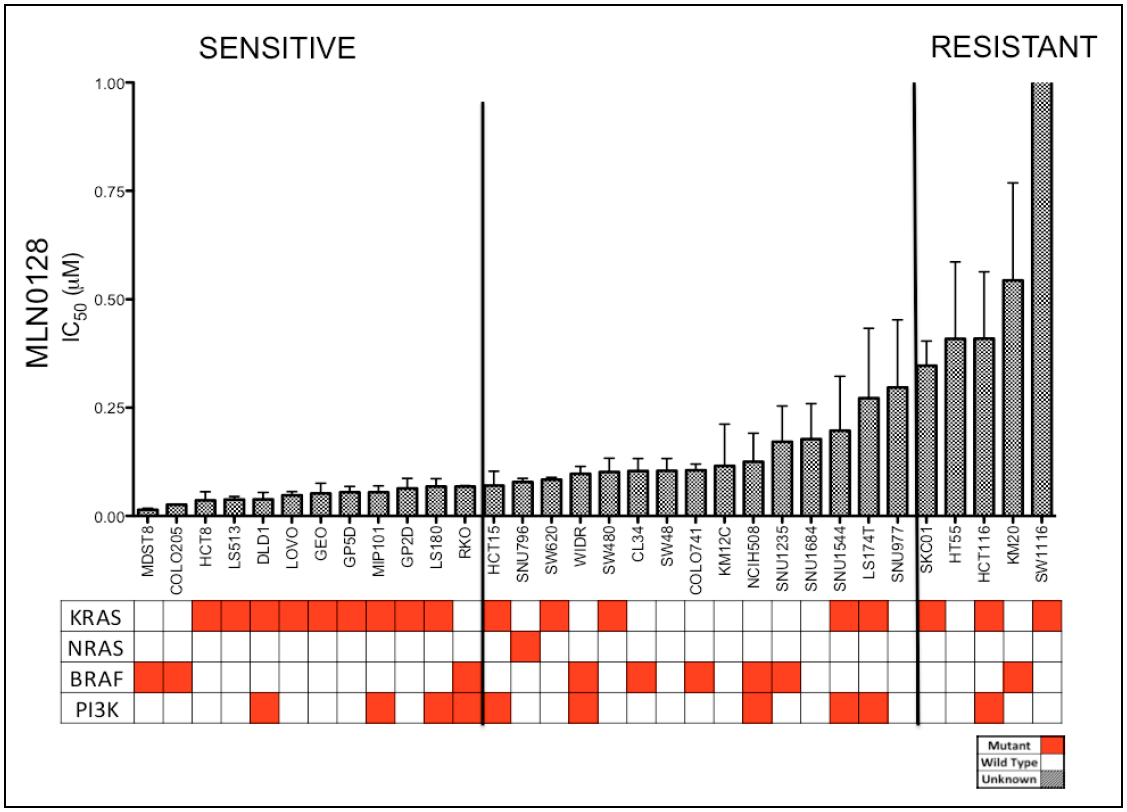


**Figure 5. Training Set for MLN0128 and their mutational status of KRAS, NRAS, BRAF and PIK3CA.**

For training the MLN0128 gene classifier, we used the cell lines as with IC50 < 0.1μM and IC50 > 0.3 μM as sensitive and resistant cell lines (**Figure 5**). There are 12 sensitive and 10 resistant cell lines selected as the training set for MLN0128 (**Table 2**).

**Table 2: MLN0128 Training Set and Genes included in the k-TSP classifier.**

| Cell Line | *LOC100505640* | *NGFR* | *LOC100506582* | *CRCT1* | *LIPC* | *MGC102966* | *TCP10L* | *PAGE2B* | *LY6G5B* | *PARD6B* | *CCNH* | *TATDN2* | *ERAL1* | *TMEM159* | **MLN0128** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COLO205 | 0.040 | 0.000 | 0.065 | 0.000 | 1.673 | 0.029 | 0.060 | 0.000 | 3.076 | 2.539 | 51.247 | 3.653 | 19.892 | 5.634 | SEN |
| DLD1 | 0.025 | 0.011 | 0.053 | 0.000 | 0.079 | 0.078 | 0.243 | 0.042 | 10.963 | 6.521 | 11.944 | 8.489 | 13.585 | 11.294 | SEN |
| GEO | 0.019 | 0.004 | 0.015 | 0.000 | 0.000 | 0.000 | 0.071 | 0.063 | 11.853 | 5.614 | 23.516 | 7.346 | 12.512 | 10.322 | SEN |
| GP2D | 0.000 | 0.000 | 0.033 | 0.000 | 0.000 | 0.000 | 0.233 | 0.000 | 3.577 | 3.252 | 18.132 | 15.839 | 27.188 | 17.368 | SEN |
| GP5D | 0.021 | 0.010 | 0.063 | 0.000 | 0.089 | 0.000 | 0.158 | 0.109 | 4.960 | 3.041 | 17.243 | 10.829 | 25.971 | 19.884 | SEN |
| HCT8 | 0.116 | 0.061 | 0.017 | 0.000 | 0.235 | 0.000 | 0.247 | 0.073 | 4.844 | 4.102 | 17.974 | 14.424 | 22.800 | 11.361 | SEN |
| LOVO | 0.048 | 0.045 | 0.055 | 0.000 | 0.427 | 0.068 | 0.076 | 0.039 | 36.403 | 2.739 | 8.740 | 6.537 | 14.300 | 6.259 | SEN |
| LS180 | 0.098 | 0.012 | 0.018 | 0.000 | 0.026 | 0.002 | 0.276 | 0.000 | 2.928 | 2.412 | 20.912 | 9.125 | 24.920 | 12.733 | SEN |
| LS513 | 0.020 | 0.000 | 0.015 | 0.000 | 0.769 | 0.002 | 0.526 | 0.066 | 4.581 | 4.315 | 16.856 | 3.094 | 12.378 | 8.481 | SEN |
| MDST8 | 0.113 | 0.050 | 0.029 | 0.000 | 0.000 | 0.000 | 0.049 | 0.000 | 2.602 | 1.219 | 18.058 | 11.586 | 16.997 | 16.221 | SEN |
| MIP101 | 0.021 | 0.000 | 0.033 | 0.000 | 0.159 | 0.052 | 0.390 | 0.103 | 7.268 | 5.356 | 13.665 | 13.134 | 15.789 | 11.417 | SEN |
| RKO | 0.025 | 0.000 | 0.054 | 0.000 | 0.014 | 0.000 | 0.053 | 0.000 | 3.508 | 1.224 | 47.250 | 6.893 | 26.532 | 16.903 | SEN |
| SKCO1 | 0.022 | 0.043 | 0.070 | 0.115 | 0.082 | 0.189 | 0.337 | 0.422 | 5.522 | 6.827 | 12.229 | 10.920 | 14.401 | 29.234 | RES |
| SNU1235 | 0.000 | 0.044 | 0.032 | 0.127 | 0.012 | 0.023 | 0.064 | 0.139 | 3.619 | 12.070 | 31.791 | 7.003 | 8.140 | 19.188 | RES |
| SNU1544 | 0.295 | 0.323 | 0.000 | 0.000 | 0.000 | 0.014 | 0.124 | 0.327 | 1.233 | 3.872 | 16.116 | 6.776 | 20.236 | 26.247 | RES |
| SNU1684 | 0.168 | 0.281 | 0.000 | 0.000 | 0.000 | 0.000 | 0.026 | 0.034 | 1.598 | 4.441 | 18.971 | 9.104 | 19.204 | 20.932 | RES |
| SNU977 | 0.124 | 0.252 | 0.000 | 0.000 | 0.013 | 0.072 | 0.097 | 0.308 | 4.297 | 14.107 | 16.303 | 3.700 | 15.585 | 44.735 | RES |
| SW1116 | 0.000 | 0.044 | 0.000 | 0.000 | 0.000 | 0.057 | 0.033 | 0.602 | 4.423 | 8.212 | 22.898 | 4.318 | 14.838 | 19.505 | RES |
| HCT116 | 0.000 | 14.375 | 0.016 | 0.081 | 0.023 | 0.076 | 0.132 | 0.387 | 3.145 | 11.278 | 22.540 | 4.625 | 8.936 | 15.008 | RES |
| HT55 | 0.000 | 0.116 | 0.000 | 0.000 | 0.000 | 0.020 | 0.025 | 0.738 | 2.939 | 4.836 | 16.358 | 9.453 | 14.643 | 15.648 | RES |
| KM20 | 0.023 | 0.224 | 0.018 | 0.563 | 0.024 | 0.071 | 0.082 | 0.108 | 2.847 | 10.627 | 20.752 | 5.524 | 13.262 | 19.212 | RES |
| LS174T | 0.111 | 0.266 | 0.000 | 0.000 | 0.000 | 0.042 | 0.188 | 0.353 | 2.559 | 3.831 | 27.872 | 4.141 | 13.322 | 17.265 | RES |

The k-TSP algorithm was applied on the training data set to generate a 7 gene pairs classifier. The decision rules for the MLN0128 classifier is defined as:

IF LOC100505640 > NGFR, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF LOC10050658 > CRCT1, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF LIPC > MGC102966, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF TCP10L > PAGE2B, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF LY6G5B > PARD6B, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF CCNH > TATDN2, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

IF ERAL1 > TMEM159, THEN Predict MLN0128 Sensitive, ELSE Predict Resistant.

As each decision rule is making a prediction, the final classifier will predict a new sample as Sensitive if 4 out of 7 rules are predicting sensitive, otherwise it will predict the new sample as resistant. The gene expressions (in FPKM values) for these 7 gene pairs were tabulated in **Table 2**.

**MLN8237 Training Set and k-TSP classifier.**

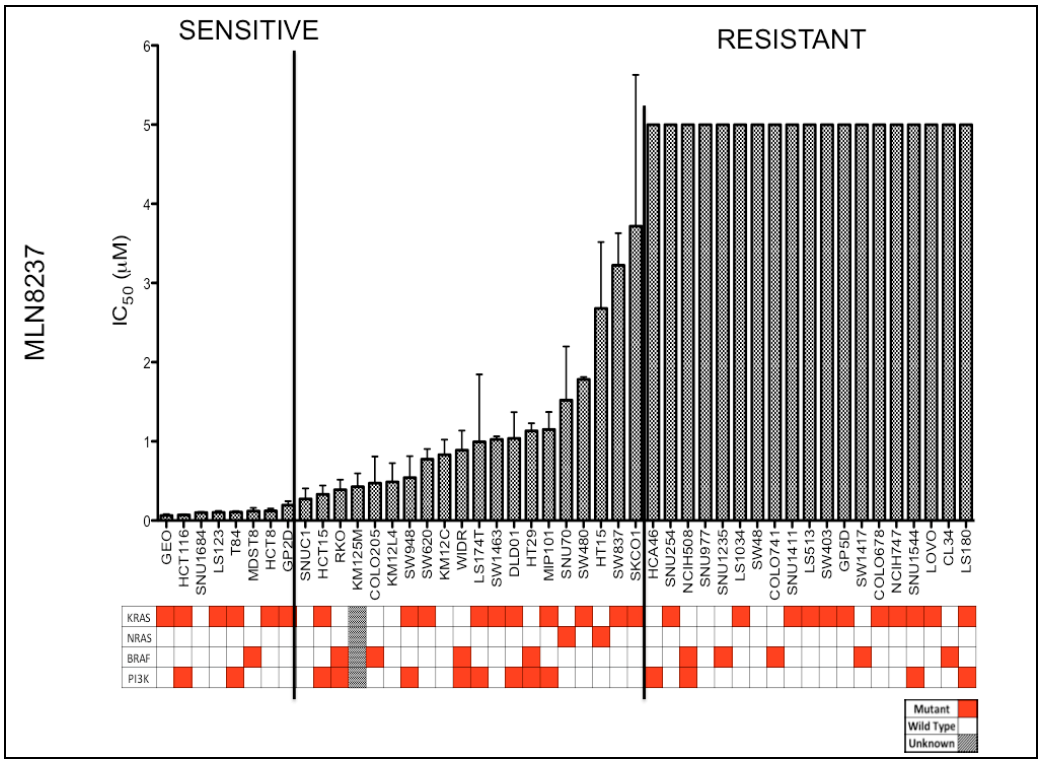


**Figure 6. Training Set for MLN8237 and their mutational status of KRAS, NRAS, BRAF and PIK3CA.**

For training the MLN8237 gene classifier, we used the cell lines as with IC50 < 0.1μM and IC50 > 5 μM as sensitive and resistant cell lines (**Figure 5**). There are 7 sensitive and 19 resistant cell lines selected as the training set for MLN8237 (**Table 2**). The k-TSP algorithm was applied on the training data set to generate a 3 gene pairs classifier. The decision rules for the MLN8237 classifier is defined as:

IF FAM177A1 > RAB24, THEN Predict MLN8237 Sensitive, ELSE Predict Resistant.

IF MEF2C > CITED1, THEN Predict MLN8237 Sensitive, ELSE Predict Resistant.

IF PRR5L > ACSS1, THEN Predict MLN8237 Sensitive, ELSE Predict Resistant.

As each decision rule is making a prediction, the final classifier will predict a new sample as Sensitive if 2 out of 3 rules are predicting sensitive, otherwise it will predict the new sample as resistant. The gene expressions (in FPKM values) for these 3 gene pairs were tabulated in **Table 3**.

**Table 3: MLN8237 Training Set and Genes included in the k-TSP classifier.**

| Cell Line | *FAM177A1* | *RAB24* | *MEF2C* | *CITED1* | *PRR5L* | *ACSS1* | **MLN8237** |
|---|---|---|---|---|---|---|---|
| GEO | 11.301 | 9.978 | 6.283 | 0.575 | 9.172 | 1.000 | SEN |
| GP2D | 18.444 | 17.555 | 1.235 | 0.214 | 9.955 | 6.167 | SEN |
| HCT116 | 10.899 | 10.678 | 9.409 | 1.742 | 0.046 | 0.029 | SEN |
| HCT8 | 14.950 | 7.733 | 7.192 | 0.356 | 1.366 | 1.216 | SEN |
| LS123 | 13.377 | 6.432 | 0.666 | 0.244 | 8.313 | 4.742 | SEN |
| MDST8 | 11.434 | 10.566 | 4.480 | 0.117 | 4.571 | 0.171 | SEN |
| SNU1684 | 17.332 | 15.333 | 6.249 | 0.139 | 16.543 | 4.371 | SEN |
| CL34 | 6.333 | 9.460 | 0.048 | 2.012 | 3.219 | 5.275 | RES |
| COLO678 | 8.538 | 13.934 | 0.062 | 0.196 | 3.246 | 14.768 | RES |
| COLO741 | 5.847 | 21.676 | 0.545 | 16.635 | 0.255 | 7.911 | RES |
| GP5D | 14.116 | 24.819 | 0.111 | 0.196 | 1.149 | 10.243 | RES |
| HCA46 | 8.366 | 8.971 | 0.287 | 0.316 | 2.541 | 3.674 | RES |
| LOVO | 4.405 | 13.726 | 1.220 | 3.876 | 3.866 | 17.594 | RES |
| LS1034 | 5.048 | 8.245 | 0.179 | 0.399 | 7.447 | 9.832 | RES |
| LS180 | 12.837 | 14.801 | 2.714 | 4.392 | 5.975 | 9.274 | RES |
| LS513 | 5.586 | 13.559 | 0.061 | 1.081 | 5.731 | 13.028 | RES |
| NCIH508 | 14.662 | 16.121 | 0.621 | 0.673 | 9.936 | 20.840 | RES |
| NCIH747 | 6.041 | 17.397 | 0.145 | 0.349 | 0.226 | 1.693 | RES |
| SNU1235 | 9.922 | 12.687 | 0.212 | 0.424 | 0.907 | 15.893 | RES |
| SNU1411 | 10.103 | 15.770 | 0.045 | 0.851 | 0.189 | 1.129 | RES |
| SNU1544 | 11.733 | 18.717 | 0.474 | 8.104 | 38.639 | 2.426 | RES |
| SNU254 | 8.656 | 20.698 | 0.000 | 9.590 | 1.113 | 2.638 | RES |
| SNU977 | 12.331 | 18.805 | 5.158 | 0.204 | 3.102 | 5.617 | RES |
| SW1417 | 6.407 | 19.553 | 0.161 | 0.285 | 8.008 | 14.742 | RES |
| SW403 | 4.360 | 11.970 | 0.010 | 0.545 | 4.419 | 17.350 | RES |
| SW48 | 12.323 | 12.525 | 0.249 | 0.842 | 1.521 | 4.808 | RES |

**ENMD2076 Training Set and k-TSP classifier.**

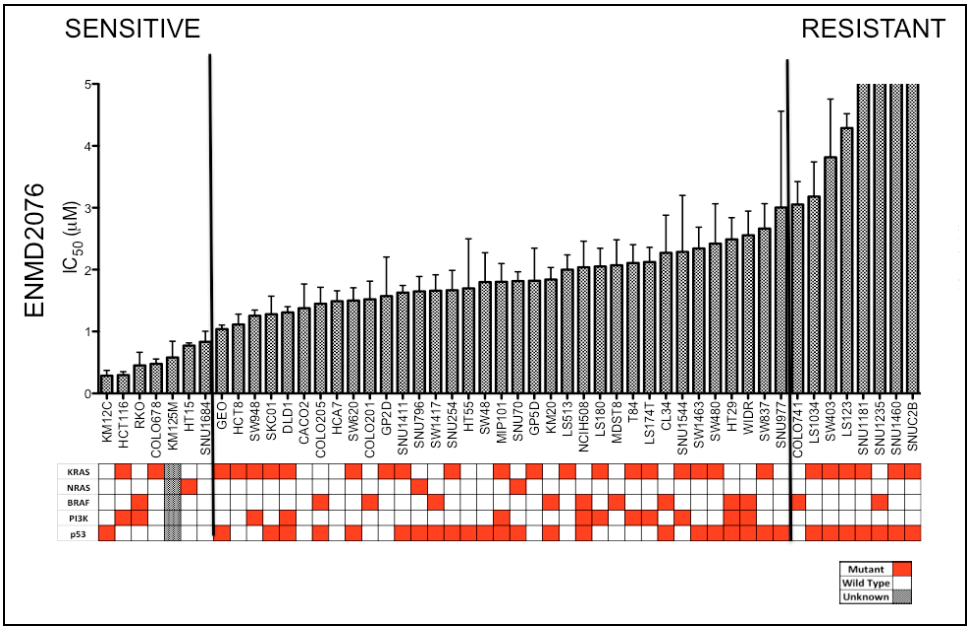


**Figure 7. Training Set for ENMD2076 and their mutational status of KRAS, NRAS, BRAF, PIK3CA and TP53.**

For training the ENMD2076 gene classifier, we used the cell lines as with IC50 < 1μM and IC50 > 3 μM as sensitive and resistant cell lines (**Figure 7**). There are 6 sensitive and 9 resistant cell lines selected as the training set for ENMD2076 (**Table 4**). The k-TSP algorithm was applied on the training data set to generate a 5 gene pairs classifier. The decision rules for the ENMD-2076 classifier is defined as:

IF AHRR > ENTPD8, THEN Predict ENMD2076 Sensitive, ELSE Predict Resistant.

IF RAMP1 > QPRT, THEN Predict ENMD2076 Sensitive, ELSE Predict Resistant.

IF GSTT1 > UBD, THEN Predict ENMD2076 Sensitive, ELSE Predict Resistant.

IF LENG9 > CDKN2B, THEN Predict ENMD2076 Sensitive, ELSE Predict Resistant.

IF IQCH > SAA2, THEN Predict ENMD2076 Sensitive, ELSE Predict Resistant.

As each decision rule is making a prediction, the final classifier will predict a new sample as Sensitive if 3 out of 5 rules are predicting sensitive, otherwise it will predict the new sample as resistant. The gene expressions (in FPKM values) for these 5 gene pairs were tabulated in **Table 4**.

**Table 4: ENMD2076 Training Set and Genes included in the k-TSP classifier.**

| **Cell Line** | *AHRR* | *ENTPD8* | *RAMP1* | *QPRT* | *GSTT1* | *UBD* | *LENG9* | *CDKN2B* | *IQCH* | *SAA2* | **ENMD2076** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| COLO678 | 2.639 | 1.365 | 0.225 | 0.079 | 25.042 | 0.000 | 4.336 | 0.000 | 0.805 | 0.392 | SEN |
| HCT116 | 2.315 | 2.171 | 17.550 | 0.343 | 0.036 | 0.015 | 4.319 | 2.512 | 1.541 | 0.186 | SEN |
| HT15 | 4.175 | 0.000 | 49.543 | 7.930 | 9.273 | 0.314 | 6.752 | 5.183 | 2.547 | 0.054 | SEN |
| KM12C | 1.746 | 0.000 | 49.181 | 0.219 | 0.222 | 0.000 | 11.725 | 2.225 | 0.300 | 0.011 | SEN |
| RKO | 0.000 | 0.000 | 3.611 | 0.167 | 13.060 | 0.017 | 2.479 | 2.173 | 0.291 | 0.000 | SEN |
| SNU1684 | 3.036 | 0.000 | 28.310 | 1.053 | 23.346 | 0.164 | 5.645 | 0.117 | 0.572 | 0.132 | SEN |
| COLO741 | 0.000 | 0.153 | 0.000 | 33.721 | 0.016 | 0.168 | 0.000 | 0.013 | 0.000 | 0.021 | RES |
| LS1034 | 1.035 | 17.435 | 0.592 | 21.661 | 0.115 | 11.079 | 0.000 | 0.663 | 0.000 | 0.417 | RES |
| LS123 | 0.000 | 3.493 | 1.840 | 6.694 | 5.238 | 10.786 | 0.000 | 11.466 | 0.000 | 18.215 | RES |
| SNU1181 | 0.000 | 0.724 | 10.870 | 102.533 | 1.111 | 9.775 | 0.000 | 16.636 | 0.000 | 1.495 | RES |
| SNU1235 | 5.629 | 6.968 | 5.753 | 13.369 | 0.037 | 0.195 | 8.450 | 17.054 | 0.879 | 17.385 | RES |
| SNU1460 | 0.000 | 1.771 | 0.551 | 59.378 | 18.272 | 238.050 | 5.799 | 8.283 | 0.000 | 9.410 | RES |
| SNU977 | 0.000 | 4.884 | 10.256 | 42.595 | 0.062 | 0.397 | 7.299 | 47.233 | 0.589 | 2.643 | RES |
| SNUC2B | 0.000 | 2.552 | 0.042 | 5.105 | 0.029 | 7.208 | 0.000 | 2.157 | 0.428 | 102.972 | RES |
| SW403 | 0.000 | 4.236 | 6.457 | 26.308 | 0.000 | 0.030 | 0.000 | 0.212 | 0.000 | 0.000 | RES |

**Aim 2.** To validate the preclinical efficacy of these classifiers against 20 independent patient-derived CRC explant models.

**Task 1: Prediction of the human CRC explants (Months 24-36, Drs. Eckhardt and Tan)**
Using the classifiers developed in Aim 1 for the three agents, we predicted the response to the drugs using baseline mRNA-seq expression obtained from CRC PDX models. **See Dr. Eckhardt's report** for the response graphs. The results for Aim 2 are provided in **Tables 5, 6, and 7**.

**Task 2: The human CRC explants will be treated with the agent and assessed for response (Months 24-36, Dr. Eckhardt).**
The classifiers predictions in **Task 1** of **Aim 2** were validated in PDX models**. See Dr. Eckhardt's report**. The results for Aim 2 are provided in **Tables 5, 6, and 7**.

**Classifier Assessment.**
We used the following three metrics to assess the performance of the classifier based on the prediction as validated by the PDX models response to the three agents.

| | | Actual PDX Response | |
|---|---|---|---|
| | | SEN | RES |
| Prediction of PDX Response | SEN | TP | FP |
| | RES | FN | TN |

**Accuracy:** 100 x (TP + TN)/(TP + FP + FN + TN)

**Sensitivity:** 100 x (TP)/(TP + FN)

**Specificity:** 100 x (TN)/(TN + FP)

**MLN0128 Prediction on the PDX models.**
Using the MLN0128 gene classifier developed in **Aim 1 (Table 2)**, we predict the PDX models response to MLN0128. As tabulated in **Table 5**, the classifier correctly predicted 13 out of 18 (72% accuracy) PDX models response to MLN0128. The sensitivity and specificity of the MLN0128 classifier are 67% and 78%, respectively. The gene pairs expressed of the MLN0128 classifier (in FPKM values) and the prediction results were tabulated in **Table 5**.

**Table 5. Prediction of the MLN0128 classifier on 20 PDX models.**

| PDX | *LOC100505640* | *NGFR* | *LOC100506580* | *CRCT1* | *LIPC* | *MGC102966* | *TCP10L* | *PAGE2B* | *LY6G5B* | *PARD6B* | *CCNH* | *TATDN2* | *ERAL1* | *TMEM159* | MLN0128 | PREDICTION CORRECT? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CRC166 | 0.000 | 0.111 | 0.118 | 0.000 | 0.000 | 0.047 | 0.244 | 0.000 | 13.876 | 9.675 | 26.764 | 15.873 | 57.188 | 11.175 | SEN | YES |
| CRC172 | 0.000 | 0.040 | 0.000 | 0.000 | 0.144 | 0.000 | 0.170 | 0.390 | 9.884 | 5.405 | 20.137 | 7.798 | 23.129 | 16.523 | SEN | YES |
| CRC026 | 0.000 | 2.040 | 0.000 | 0.067 | 0.723 | 0.094 | 0.097 | 0.173 | 5.143 | 2.884 | 10.637 | 6.016 | 12.062 | 9.424 | SEN | YES |
| CRC098 | 0.000 | 0.005 | 0.000 | 0.000 | 0.037 | 0.009 | 0.054 | 0.036 | 6.237 | 3.615 | 15.420 | 7.231 | 16.506 | 16.050 | SEN | YES |
| CRC040 | 0.000 | 0.014 | 0.000 | 0.101 | 0.000 | 0.006 | 0.017 | 0.000 | 4.039 | 2.986 | 15.687 | 15.123 | 14.106 | 23.200 | SEN | NO |
| CRC125 | 0.000 | 0.434 | 0.000 | 0.000 | 0.122 | 0.011 | 0.017 | 0.130 | 5.571 | 5.268 | 13.211 | 6.857 | 25.613 | 19.331 | SEN | YES |
| CRC102 | 0.000 | 0.016 | 0.000 | 0.000 | 0.000 | 0.000 | 0.019 | 0.000 | 9.101 | 4.304 | 22.971 | 20.872 | 20.684 | 23.670 | SEN | NO |
| CRC027 | 0.052 | 0.031 | 0.000 | 0.000 | 1.910 | 0.000 | 0.000 | 0.249 | 6.658 | 12.186 | 30.864 | 11.350 | 12.639 | 6.993 | SEN | YES |
| CRC106 | 0.000 | 0.015 | 0.000 | 0.176 | 0.000 | 0.005 | 0.094 | 0.000 | 8.092 | 4.548 | 4.989 | 7.404 | 10.133 | 25.893 | RES | YES |
| CRC042 | 0.025 | 0.000 | 0.000 | 0.105 | 0.013 | 0.030 | 0.196 | 0.038 | 7.426 | 6.271 | 15.249 | 8.379 | 14.735 | 12.755 | RES | NO |
| CRC036 | 0.000 | 0.575 | 0.000 | 0.190 | 0.146 | 0.024 | 0.066 | 0.048 | 2.273 | 5.200 | 16.478 | 8.665 | 22.078 | 19.827 | RES | NO |
| CRC021 | 0.000 | 0.027 | 0.000 | 0.148 | 0.000 | 0.086 | 0.043 | 0.048 | 1.483 | 1.275 | 7.652 | 2.372 | 4.731 | 3.842 | RES | YES |
| CRC147 | 0.270 | 3.448 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.692 | 5.748 | 7.019 | 14.761 | 10.008 | 60.230 | 18.839 | RES | YES |
| CRC001 | 0.000 | 0.004 | 0.000 | 0.000 | 0.000 | 0.012 | 0.061 | 0.252 | 5.122 | 3.348 | 11.328 | 5.495 | 13.426 | 30.318 | RES | YES |
| CRC007 | 0.000 | 0.006 | 0.000 | 0.059 | 0.000 | 0.049 | 0.031 | 0.000 | 4.575 | 2.163 | 7.026 | 3.946 | 11.235 | 10.103 | RES | NO |
| CRC006 | 0.000 | 0.156 | 0.000 | 0.000 | 0.016 | 0.108 | 0.007 | 0.422 | 6.261 | 3.192 | 10.014 | 5.903 | 11.889 | 16.279 | RES | YES |
| CRC108 | 0.039 | 0.015 | 0.000 | 0.040 | 0.022 | 0.002 | 5.324 | 0.031 | 4.118 | 4.797 | 10.041 | 11.946 | 19.584 | 22.569 | RES | YES |
| CRC162 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 2.122 | 3.190 | 28.034 | 16.219 | 32.930 | 40.136 | RES | YES |

**MLN8237 Prediction on the PDX models.**
Using the MLN8237 gene classifier developed in **Aim 1 (Table 5)**, we predict the PDX models response to MLN8237. As tabulated in **Table 6**, the classifier correctly predicted 16 out of 21 (62% accuracy) PDX models response to MLN8237. The sensitivity and specificity of the MLN8237 classifier are 50% and 67%, respectively. The gene pairs expressed of the MLN8237 classifier (in FPKM values) and the prediction results were tabulated in **Table 6**.

**Table 6. Prediction of the MLN8237 classifier on PDX models**

| PDX | *FAM177A1* | *RAB24* | *MEF2C* | *CITED1* | *PRR5L* | *ACSS1* | MLN8237 | PREDICTION CORRECT? |
|---|---|---|---|---|---|---|---|---|
| CRC162 | 38.875 | 16.435 | 0.245 | 3.024 | 11.037 | 36.653 | SEN | NO |
| CRC108 | 24.466 | 5.138 | 0.044 | 3.041 | 9.980 | 21.717 | SEN | NO |
| CRC166 | 74.510 | 15.669 | 1.919 | 1.412 | 10.465 | 25.957 | SEN | YES |
| CRC098 | 26.058 | 11.471 | 1.145 | 1.785 | 14.008 | 26.989 | SEN | NO |
| CRC020 | 3.597 | 18.322 | 0.438 | 0.671 | 0.947 | 8.043 | SEN | NO |
| CRC040 | 9.636 | 7.912 | 0.294 | 2.426 | 7.204 | 6.845 | SEN | YES |
| CRC034 | 40.024 | 11.705 | 2.050 | 1.144 | 10.052 | 24.686 | SEN | YES |
| CRC181 | 21.961 | 14.976 | 0.015 | 6.820 | 2.836 | 30.185 | SEN | NO |
| CRC026 | 8.800 | 11.382 | 0.576 | 3.229 | 8.472 | 3.931 | RES | YES |
| CRC007 | 9.466 | 16.808 | 0.052 | 2.089 | 3.380 | 9.953 | RES | YES |
| CRC010 | 6.285 | 16.472 | 0.449 | 0.541 | 9.073 | 32.108 | RES | YES |
| CRC125 | 10.024 | 10.722 | 0.157 | 2.960 | 9.280 | 24.704 | RES | YES |
| CRC006 | 11.592 | 9.927 | 1.901 | 52.841 | 16.008 | 10.418 | RES | NO |
| CRC102 | 13.173 | 15.401 | 0.030 | 0.688 | 10.529 | 21.972 | RES | YES |
| CRC001 | 9.876 | 11.437 | 0.051 | 3.218 | 3.499 | 15.333 | RES | YES |
| CRC036 | 11.222 | 9.102 | 1.436 | 0.309 | 8.764 | 11.110 | RES | NO |
| CRC021 | 3.471 | 5.122 | 1.793 | 0.034 | 2.190 | 1.607 | RES | NO |
| CRC042 | 21.217 | 12.092 | 0.047 | 2.773 | 3.864 | 12.081 | RES | YES |
| CRC047 | 9.313 | 7.323 | 0.143 | 0.194 | 5.137 | 6.272 | RES | YES |
| CRC027 | 19.840 | 19.034 | 0.024 | 0.593 | 7.479 | 9.878 | RES | YES |
| CRC012 | 5.049 | 10.572 | 0.011 | 1.530 | 3.252 | 9.244 | RES | YES |

**ENMD2076 Prediction on the PDX models.**
Using the ENMD2076 gene classifier developed in **Aim 1 (Table 4)**, we predict the PDX models response to ENMD2076. As tabulated in **Table 7**, the classifier correctly predicted 16 out of 20 (80% accuracy) PDX models response to ENMD-2076. The sensitivity and specificity of the ENMD2076 classifier are 100% and 73%, respectively. The gene pairs expressed of the ENMD2076 classifier (in FPKM values) and the prediction results were tabulated in **Table 7**.

**Table 7. Prediction of the ENMD2076 classifier on PDX models.**

| PDX | *AHRR* | *ENTPD8* | *RAMP1* | *QPRT* | *GSTT1* | *UBD* | *LENG9* | *CDKN2B* | *IQCH* | *SAA2* | ENMD2076 | PREDICTION CORRECT? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CRC181 | 152.972 | 2.180 | 237.597 | 0.048 | 28.729 | 0.000 | 8.495 | 2.387 | 0.193 | 0.000 | SEN | YES |
| CRC021 | 0.000 | 2.381 | 0.285 | 14.407 | 7.760 | 1.658 | 0.000 | 0.264 | 0.073 | 0.568 | SEN | NO |
| CRC098 | 181.412 | 9.241 | 26.886 | 25.026 | 0.000 | 411.709 | 9.700 | 5.819 | 0.110 | 0.814 | SEN | YES |
| CRC006 | 0.000 | 10.906 | 55.682 | 3.758 | 17.619 | 0.845 | 0.000 | 2.569 | 0.000 | 0.240 | SEN | NO |
| CRC166 | 90.684 | 13.744 | 3.593 | 90.608 | 31.575 | 2.556 | 7.244 | 0.780 | 2.161 | 0.000 | SEN | YES |
| CRC026 | 0.000 | 8.171 | 28.296 | 21.990 | 39.966 | 27.325 | 0.000 | 0.516 | 0.000 | 2.399 | SEN | NO |
| CRC042 | 0.000 | 4.680 | 1.483 | 18.534 | 16.942 | 0.421 | 5.400 | 1.621 | 0.276 | 0.063 | SEN | YES |
| CRC162 | 205.872 | 3.010 | 2.130 | 12.272 | 0.000 | 0.436 | 7.487 | 3.685 | 2.595 | 22.077 | SEN | NO |
| CRC106 | 0.000 | 16.812 | 55.490 | 1.253 | 13.592 | 5.428 | 10.839 | 2.462 | 0.495 | 0.146 | SEN | YES |
| CRC027 | 0.000 | 0.000 | 2.107 | 128.320 | 47.267 | 2.849 | 0.000 | 0.353 | 0.000 | 0.000 | RES | YES |
| CRC036 | 0.000 | 5.194 | 3.708 | 72.932 | 22.964 | 57.259 | 0.000 | 3.950 | 0.333 | 0.929 | RES | YES |
| CRC047 | 0.000 | 1.963 | 0.056 | 25.347 | 0.079 | 17.509 | 0.000 | 1.854 | 0.309 | 1.538 | RES | YES |
| CRC108 | 0.000 | 3.037 | 43.469 | 45.804 | 17.215 | 1.134 | 0.000 | 3.428 | 1.228 | 0.153 | RES | YES |
| CRC102 | 0.000 | 2.010 | 1.820 | 33.681 | 37.107 | 15.856 | 0.000 | 2.473 | 0.431 | 0.872 | RES | YES |
| CRC125 | 0.000 | 2.541 | 132.330 | 19.687 | 11.618 | 6.343 | 0.000 | 6.930 | 0.718 | 4.046 | RES | YES |
| CRC001 | 0.000 | 4.913 | 9.864 | 10.686 | 14.634 | 9.857 | 0.000 | 1.294 | 0.000 | 0.032 | RES | YES |
| CRC034 | 0.000 | 0.000 | 32.664 | 38.509 | 0.131 | 6.350 | 0.000 | 3.292 | 0.764 | 0.166 | RES | YES |
| CRC007 | 0.000 | 0.000 | 1.900 | 10.772 | 20.378 | 0.074 | 5.403 | 1.490 | 0.000 | 0.000 | RES | YES |
| CRC040 | 0.000 | 0.000 | 0.926 | 19.569 | 0.071 | 1.305 | 0.000 | 3.316 | 1.209 | 10.112 | RES | YES |
| CRC012 | 0.000 | 7.284 | 2.551 | 17.402 | 10.443 | 1.170 | 4.473 | 0.324 | 0.000 | 0.302 | RES | YES |

**KEY RESEARCH ACCOMPLISHMENTS:**

1. Completed in vitro screening on a large panel of CRC cell lines to determine the activity of three novel anti-cancer agents
2. Completed in vivo screening on CRC patient-derived xenografts
3. Completed baseline gene expression profiling of CRC cell lines and patient-derived tumor explants by high-throughput RNA-sequencing approach
4. Analyzed the RNA-seq data with bioinformatics pipeline
5. Developed initial predictive classifiers for the three novel anti-cancer agents

**REPORTABLE OUTCOMES**: Based on the data obtained we have published two manuscripts and three additional manuscripts are in preparation.

1. Davis SL, Robertson KM, Pitts TM, Tentler JJ, Bradshaw-Pierce EL, Klauck PJ, Bagby SM, Hyatt SL, Selby HM, Spreafico A, Ecsedy JA, Arcaroli JJ, Messersmith WA, Tan AC, Eckhardt SG. (2015). Combined inhibition of MEK and Aurora A kinase in KRAS/PIK3CA double-mutant colorectal cancer models. *Front Pharmacol*. 6:120. [PMID: 26136684]. [PMCID: PMC4468631].

**2.** Christopher H. Lieu, Patrick K. Henthorn, John J. Tentler, Aik-Choon Tan, Anna Spreafico, Heather M. Selby, Stacey M. Bagby, Peter J. Klauck, John J. Arcaroli, Wells A. Messersmith, Todd M. Pitts , S. Gail Eckhardt. Antitumor Activity of the Potent MEK Inhibitor, TAK733, Against Colorectal Cancer Cell Lines and Patient Derived Xenografts. *Oncotarget*. 6(33):34561-34572. [PMID: 26439693].

**3.** Todd M Pitts, Erica L Bradshaw-Pierce, Stacey M Bagby, Stephanie L Hyatt, Heather M Selby, Anna Spreafico, John J Tentler, Kelly McPhillips, Peter J Klauck, Anna Capasso, Aik Choon Tan, John J Arcaroli, Alicia Purkey, Wells A Messersmith, Jeffery A Ecsedy, S Gail Eckhardt. Antitumor Activity of the Aurora A Selective Kinase Inhibitor, Alisertib, Against Preclinical Models of Colorectal Cancer. *In Preparation.*

**4.** Anna Capasso, Todd M Pitts, John J Tentler, Peter J Klauck, Anna Capasso, Aik Choon Tan, John J Arcaroli, Alicia Purkey, Wells A Messersmith, S Gail Eckhardt. Dual Compartmental Targeting of Cell Cycle and Angiogenic Kinases in Colorectal Cancer Models by ENMD2076. *In Preparation.*

**5.** Peter J Klauck, Todd M Pitts, Aik Choon Tan, John J Tentler, John J Arcaroli, Alicia Purkey, Wells A Messersmith, S Gail Eckhardt. Antitumor Activity of the Polo-Like Kinase 1 Inhibitor, TAK960, Against Preclinical Models of Colorectal Cancer. *In Preparation*.

**Other publications, conference papers, and presentations (See appendices):**

1. Tan AC, Britt BW, Astling DP, Leong S, Lieu C, Tentler JJ, Pitts TM, Arcaroli JJ, Messersmith WA, Eckhardt SG. (2012). Validation of Preclinical Colorectal Cancer Models Against TCGA Data for Pathway Analysis and Predictive Biomarker Discovery. (Presented in the EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, Dublin, Ireland.)

2. T. Pitts[1], K.L. McPhillips[1], H.M. Selby[1], A. Spreafico[1], S.M. Bagby[1], B.C. Britt[1], J.J. Tentler[1], A.C. Tan[1], K. Kuida[2], S.G. Eckhardt[1]. Antitumor Activity of the Polo-like Kinase (PLK) Inhibitor, TAK-960, Alone and in Combination with Standard Agents Against KRAS WT and MT Colorectal Cancer (CRC) Models 1 University of Colorado, Medical Oncology, Aurora CO, USA; 2Millennium: The Takeda Oncology Company, Translational Medicine, Cambridge MA, USA (European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 172-173, Poster 562)

3. J Tentler, SM Bagby, AC Tan, TM Pitts, HM Selby, KL McPhillips, SG Eckhardt, S Leong

Molecular Markers of Sensitivity to the Aurora and Angiogenic Kinase Inhibitor ENMD-2076 in Human Colorectal Cancer (CRC) Models. University of Colorado, Medical Oncology, Aurora CO, USA. (European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 78, Poster 255)

4. TM Pitts[1], KL McPhillips[1], HM Selby[1], A Spreafico[1], SM Bagby[1], BC Britt[1], AC Tan[1], JJ Tentler[1], JA Ecsedy[2], SG Eckhardt[1]. In Vitro and in Vivo Antitumor Activity of the Investigational Aurora A Selective Kinase Inhibitor MLN8237 Alone and in Combination with Standard Agents Against CRC Models. [1]University of Colorado, Medical Oncology, Aurora CO, USA. [2]Millennium: The Takeda Oncology Company, Translational Medicine, Cambridge MA, USA. (European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 78, Poster 254)

**CONCLUSIONS:** The overall impact of the work performed in this proposal is largely related to the fact that we were able to identify several novel agents that are active against CRC *in vitro* and *in vivo.* The scientific climate has changed somewhat in the last 5 years since there is much greater focus on rational combinations and combinations with immunotherapy in oncology, so that the use of classifiers for single-agent treatment currently has limited clinical application. Nonetheless, we are moving forward with these agents in CRC but with an eye towards combination strategies, and are developing humanized mouse models of our CRC PDX so that we can test combination strategies with these novel agents and immunotherapy. Ideally, we hope to integrate the biomarker data obtained in this proposal in order to select rational combinations for patients. We have completed all of the **Tasks in Aim 1 and Aim 2.** In the future, we are using these data to predict rational combinations of these novel anti-cancer agents in CRC.

**REFERENCES:**


Diamond JR, Eckhardt SG, Tan AC, Newton TP, Selby HM, Brunkow KL, Kachaeva MI, Varella-Garcia M, Pitts TM, Bray MR, Fletcher GC, Tentler JJ. (2013). Predictive Biomarkers of Sensitivity to the Aurora and Angiogenic Kinase Inhibitor ENMD-2076 in Preclinical Breast Cancer Models. Clinical Cancer Research. 19(1): 19:291-303.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297-1303.

Pitts TM, Tan AC, Kulikowski GN, Tentler JJ, Brown AM, Flanigan SA, Leong S, Coldren CD, Hirsch FR, Varella-Garcia M, Korch C, Eckhardt SG. (2010). Development of an integrated genomic classifier for a novel agent in colorectal cancer: approach to individualized therapy in early development. Clin Cancer Res. 16(12):3193-3204.

Skehan P, Storeng R, Scudiero D, Monks A, McMahon J, Vistica D, Warren JT, Bokesch H, Kenney S, Boyd MR. (1990). New colorimetric cytotoxicity assay for anticancer-drug screening. J Natl Cancer Inst. 82(13):1107-1112.

Tan AC, Naiman DQ, Xu L, Winslow RL, Geman D. (2005) Simple decision rules for classifying human cancers from gene expression profiles. Bioinformatics. 21(20):3896-3904.

Trapnell, C., Pachter, L., and Salzberg, S.L. 2009. TopHat: discovering splice junctions with RNAseq. Bioinformatics 25:1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. 2010. Transcript assembly and quantification by RNAseq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511-515.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38(16):e164.

**Appendices:**
**Abstract** Presented in the EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, Dublin, Ireland.

**Validation of Preclinical Colorectal Cancer Models Against TCGA Data for Pathway Analysis and Predictive Biomarker Discovery**

A. Tan[1], B. Britt[1], D. Astling[1], S. Leong[1], C. Lieu[1], J. Tentler[1], T. Pitts[1], J. Arcaroli[1], W. Messersmith[1], S. Eckhardt[1].
[1]University of Colorado Anschutz Medical Campus, Medical Oncology/Medicine, Aurora CO, USA

**Background:** Preclinical models such as cancer cell lines and patient- derived tumor xenografts (PDTX) have been widely used in predictive biomarker development and pathway modeling in cancer research. However, it has not been clear to what extent these preclinical models reflect the molecular heterogeneity observed in clinical samples, while initiatives such as the TCGA provide an opportunity for comparison and validation.

**Methods:** We performed massively parallel mRNA sequencing (RNA-seq) on 25 PDTX and 60 CRC cell lines using the Illumina HiSeq2000 platform to characterize the transcriptome of these preclinical models. On average, 40 million single-end 100bp sequencing reads per sample were obtained. The RNA-seq reads were mapped against the human genome using Tophat (version 1.3.2). On average, 80% of the reads aligned to the human genome. Cufflinks (version 1.3.0) was used to assemble the transcripts using the RefSeq annotation as the guide. Gene-level expression was estimated by FPKM (fragments per kilobase of exon per million fragments mapped). We performed pathway analysis using PARADIGM. RNA-seq of 244 CRC patient tumors were downloaded from the TCGA website. Following rank-normalized, mean centered data normalization, hierarchical clustering was performed on the samples using gene-centric and pathway- centric approaches.

**Results:** To determine whether the preclinical models were representative of the variability observed in expression profiles from clinical samples, we compared RNA-seq gene expression data of the 25 PDTX and 60 CRC cell lines with 244 TCGA CRC patient tumors. From the unsupervised hierarchical clustering approach, CRC cell lines and PDTX clustered together with TCGA patient tumors. We also performed unsupervised hierarchical clustering based on PARADIGM inferred gene sets. In the pathway clustering analysis, the preclinical CRC models also clustered together with TCGA patient samples. Within each cluster, CRC preclinical models do response to particular class of targeted therapy, suggesting potential treatment strategies for the diverse CRC patient samples.

**Conclusions:** In this study, we performed a systematic comparison of our CRC preclinical models and TCGA patient samples using next-generation sequencing data. Clustering analysis indicates that our preclinical models are representative of all CRC patient clusters identified in TCGA database. These results indicate that these CRC preclinical models are representative of actual patient samples and may be useful in early drug development and predictive biomarker discovery.

(European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 81, Poster 263)

**Antitumor Activity of the Polo-like Kinase (PLK) Inhibitor, TAK-960, Alone and in Combination with Standard Agents Against KRAS WT and MT Colorectal Cancer (CRC) Models**

T. Pitts[1], K.L. McPhillips[1], H.M. Selby[1], A. Spreafico[1], S.M. Bagby[1], B.C. Britt[1], J.J. Tentler[1], A.C. Tan[1], K. Kuida[2], S.G. Eckhardt[1].
1 University of Colorado, Medical Oncology, Aurora CO, USA;
2Millennium: The Takeda Oncology Company, Translational Medicine, Cambridge MA, USA

**Background:** Polo-like kinases (PLKs) are serine-threonine kinases that are involved in several processes of cell division including chromosomal segregation, spindle formation, and cytokinesis. PLKs, specifically PLK-1, are highly expressed in cells and tissues with high mitotic indices such as cancer, and are overexpressed in head and neck, lung, breast and colon malignancies, among others. In this preclinical study we assessed the antitumor effects of the novel Plk inhibitor, TAK-960, against CRC models, including cell lines and patient-derived xenografts.

**Methods:** The anti-proliferative effects of TAK-960 as a single agent and in combination with irinotecan (SN38) or cetuximab were assessed using an assay that measures DNA content (CyQUANT). Synergy was calculated using Calcusyn software while evaluation of downstream effector molecules and apoptosis was assessed by immunoblotting. Patient-derived CRC xenografts were implanted into athymic nude mice and tumor growth inhibition (TGI) was evaluated following treatment with TAK-960 alone or in combination with standard agents (irinotecan or cetuximab).

**Results:** CRC cell lines were quite sensitive to TAK-960 with IC50 values ranging from 0.007 to 1 umol/L. While no synergy was observed in the KRAS WT CRC cell lines in the cetuximab combination groups, additivity to mild synergy was observed in the KRAS MT CRC cell lines exposed to the SN38 combination. Modulation of down stream effector molecules was observed following exposure to TAK-960, including pHistone H3 and p73. Interestingly, against patient-derived xenograft models, synergy was difficult to assess in the KRAS WT models due to the exquisite sensitivity to cetuximab, while some of the KRAS MT xenografts did demonstrate TGI in the irinotecan combination groups that was supra-additive.

**Conclusion:** The PLK inhibitor TAK-960 demonstrated robust single-agent anti-proliferative effects against CRC cell lines in vitro, whereas synergy was not observed when combined with cetuximab or SN38. However, there were supra-additive effects noted in several patient-derived KRAS MT xenografts treated with TAK-960 and irinotecan, supporting the evaluation of this regimen in this patient population with limited therapeutic options.

(European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 172-173, Poster 562)

**TAK-733, an Investigational Novel MEK Inhibitor, Suppresses Colorectal Cancer (CRC) Tumor Growth in Biomarker Positive Patient-derived Human Tumor Explants**

CH Lieu, JL Tentler, AC Tan, TM Pitts, A Spreafico, HM Selby, KL McPhillips, SM Bagby, SG Eckhardt
University of Colorado, Medical Oncology, Aurora CO, USA

**Background:** CRC is a significant cause of cancer mortality, and new therapies are needed for patients with advanced disease. TAK-733 is a highly potent and selective investigational novel MEK allosteric site inhibitor.

**Materials and Methods:** In a preclinical study of TAK-733, a panel of CRC cell lines was exposed to varying concentrations of TAK-733 for 72 hours followed by sulforhodamine B assay. Cell lines were segregated into sensitive (IC50 □ 0.5 mM) or resistant (IC50 > 0.5mM). Twenty patient-derived human tumor explants grown in vivo as xenografts were then treated with TAK-733. Tumor growth inhibition (TGI) was measured to determine the sensitivity of the CRC explants to TAK-733. A sensitive explant was defined by a TGI□80%. Linear regression was used to examine the predictive effects of genotype on the TGI of explants.

**Results:** Fifty-four CRC cell lines were exposed to TAK-733, and 42 cell lines were found to be sensitive across a broad range of mutations within these cell lines. Eighty-two percent of the cell lines within the sensitive subset were BRAF or KRAS mutant, and 80% of the cell lines within the sensitive subset were PIK3CA WT. The predictability of these mutations is limited, because a majority (7/12) of the insensitive cell lines also contained mutations in BRAF and KRAS. Twenty patient-derived human tumor CRC explants were then treated with TAK-733. In total, 15 primary human tumor explants were found to be sensitive to TAK-733 (TGI □ 80%), including 9 primary human tumor explants exhibiting tumor regression (TGI >100%). Explants with a BRAF/KRAS mutant and PIK3CA wild-type genotype demonstrated increased sensitivity to TAK-733 with a median TGI of 106%. Published MEK-response gene signatures also correlated with response to TAK-733.

**Conclusions:** TAK-733 demonstrates robust antitumor activity against CRC cell lines and patient-derived tumor explants. There was a trend towards higher sensitivity to TAK-733 in tumors that were BRAF/KRAS mutant and PIK3CA wild-type. There was also a trend towards sensitivity to TAK-733 in tumors with published MEK-response gene signatures. This data may provide a potential patient selection strategy for future clinical trials in patients with metastatic CRC.

(European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 119-120, Poster 393)

**Molecular Markers of Sensitivity to the Aurora and Angiogenic Kinase Inhibitor ENMD-2076 in Human Colorectal Cancer (CRC) Models**

J Tentler, SM Bagby, AC Tan, TM Pitts, HM Selby, KL McPhillips, SG Eckhardt, S Leong
University of Colorado, Medical Oncology, Aurora CO, USA

**Background:** ENMD-2076 is an orally bioavailable small molecule currently in clinical development that is an inhibitor of Aurora kinase A, as well as angiogenic kinases VEGFR2 and PDGFRa. The purpose of this study was to use gene set enrichment analysis (GSEA) and RNA-seq data from preclinical models of CRC to develop predictive markers of sensitivity to ENMD-2076.

**Methods:** To determine sensitivity (S) or resistance (R), a panel of 52 CRC cell lines was exposed to increasing doses of ENMD-2076 and proliferation was measured by the sulforhodamine B method. For in vivo studies, athymic nude mice were injected subcutaneously with 3mm3 sections of patient-derived CRC tumor explants (PDTX). When tumors reached a volume of ~150 mm3 , mice were randomized into vehicle and ENMD-2076 (200mg/kg) groups; n=5 per group. Vehicle or drug was administered qd for 30 days by oral gavage with tumor volume measurements taken every 3 days. High-throughput mRNA sequencing (RNA-seq) of CRC cell lines and PDTX models was obtained using the Illumina HiSeq2000. On average, 40 million single-end 100bp sequencing reads per sample were obtained. The RNA-seq reads were mapped against the human genome using Tophat (version 1.3.2). On average, 80% of the reads aligned to the human genome. Cufflinks (version 1.3.0) was used to assemble the transcripts using the RefSeq annotation as the guide. For GSEA, pathways were obtained from KEGG and AMBION databases as gene sets. Enriched pathways were identified by running GSEA using 1000 permutations. Predictive biomarkers for ENMD-2076 sensitivity were derived from the RNA-seq data using the k-TSP learning algorithm.

**Results:** To determine the genes and pathways correlated with ENMD- 2076 responsiveness, GSEA was performed comparing baseline gene expression profiles of eleven S (IC50 □ 1mM) and five R (IC50 □ 5mM) cell lines. Six pathways were enriched in the S lines ($p < 0.01$) and 28 pathways were enriched in the R lines ($p < 0.01$). Among the top enriched pathways in the R lines were cytokine-related pathways, chemokine signaling pathways, JAK/STAT and PI3K signaling pathways. These results point to potential ra- tional combination studies with ENMD-2076 in CRC resistant cell lines. For the predictive biomarker development strategy, the k-TSP algorithm was trained on the RNA-seq data from the S and R cell lines. Gene pair classifiers were then derived and tested on the RNA-seq of ten CRC PDTX tumor models. Among the ten PDTX models, nine had a TGI <50% and were predicted S while one of the explants had a TGI >150% was predicted as R.
**Conclusions:** The results of this study indicate that it is possible to derive predictive biomarkers from CRC cell lines and predict sensitivity on CRC PDTX models. Further refinement of this classifier by including mutational data will greatly improve the robustness of these predictive biomarkers.

(European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 78, Poster 255)

**In Vitro and in Vivo Antitumor Activity of the Investigational Aurora A Selective Kinase Inhibitor MLN8237 Alone and in Combination with Standard Agents Against CRC Models**

TM Pitts[1], KL McPhillips[1], HM Selby[1], A Spreafico[1], SM Bagby[1], BC Britt[1], AC Tan[1], JJ Tentler[1], JA Ecsedy[2], SG Eckhardt[1]
[1]University of Colorado, Medical Oncology, Aurora CO, USA
[2]Millennium: The Takeda Oncology Company, Translational Medicine, Cambridge MA, USA

**Background:** The Aurora kinases are a family of serine/threonine kinases comprised of Aurora A, B, and C which execute critical steps in mitotic and meiotic progression. MLN8237 is an investigational Aurora A selective inhibitor that has demonstrated activity against a wide variety of tumor types in vitro and in vivo, including CRC. In this study the activity of MLN8237 alone and in combination with irinotecan or cetuximab was assessed in CRC cell lines and patient-derived tumor xenografts (PDTXs).

**Methods:** A panel of 55 CRC cell lines were exposed to increasing concentrations of MLN8237, alone or in combination with SN38, and assessed for proliferation by quantifying DNA content using a CyQUANT assay. Synergy was determined in the combinations using Calcusyn software, while downstream effector molecules and apoptosis were assessed by standard immunoblotting methods. For the in vivo studies, patient-derived CRC xenografts were implanted into athymic nude mice and tumor growth inhibition was evaluated following treatment with MLN8237 as single agent or in combination with irinotecan or cetuximab.

**Results:** Colon cancer cell lines demonstrated varying sensitivity to MLN8237 with IC50 values ranging from 0.08 to >5umol/L. Synergy to additivity was observed in several KRAS mutant CRC cell lines treated with MLN8237 and SN38 (CI=0.1−6.0). Following exposure to MLN8237 we observed an increase in pHistone H3 showing that MLN8237 was modulating its target. No remarkable combination effects of MLN8237 with cetuximab in KRAS WT PDTX was observed due to exquisite sensitivity to single agent cetuximab. Several KRAS MU PDTX did exhibit supra-additivity to MLN8237 and irinotecan combined, consistent with the beneficialcombinationobservedinvitrowithSN38.Analysisofdownstream effectors and markers of proliferation and apoptosis is ongoing.
**Conclusion:** MLN8237 demonstrated anti-proliferative effects against CRC cell lines with synergy observed in combination with SN38 in vitro. Moreover, in the PDTX models greater tumor growth inhibition was observed in several of the KRAS mutant xenografts treated with the combination of MLN8237 and irinotecan, indicating a potential clinical development strategy for the agent in KRAS MU CRC, where therapeutic options are limited.

(European Journal of Cancer, Volume 48, Supplement 6, November 2012, Pages 78, Poster 254)